GENETIC SUSCEPTIBILITY TO CHRONIC INFLAMMATORY INTESTINAL DISEASES AND RISK OF PANCREATIC DUCTAL ADENOCARCINOMA: A PATHWAY-BASED ANALYSIS OF GENOMIC-WIDE ASSOCIATION STUDIES

by

Fangcheng Yuan

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science

Baltimore, Maryland

April 2018

Abstract

Background: Chronic inflammation is known to play a role in pancreatic carcinogenesis. Registry-based epidemiologic studies suggest associations between chronic inflammatory intestinal diseases and pancreatic ductal adenocarcinoma (PDAC).

Methods: We examined the association between genomic regions (20 kb upstream and 20 kb downstream) surrounding germline variants for Crohn's disease, ulcerative colitis, inflammatory bowel disease (Crohn's disease and ulcerative colitis combined) and celiac disease identified in published genome-wide association studies (GWAS) and PDAC risk in 8384 cases and 11955 controls of European descent using summary statistic GWAS data from the Pancreatic Cancer Cohort Consortium (PanScan) and the Pancreatic Cancer Case Control Consortium (PanC4). We employed the summary adaptive rank truncated product (sARTP) method to test the overall association of the combined genomic regions for each respective disease.

Results: Categorization of the genomic regions for ulcerative colitis, Crohn's disease, and inflammatory bowel disease were associated with PDAC at *P*-values < 0.05 (0.0030, 0.037, and 0.0018, respectively). After excluding the regions around the previously identified GWAS loci for PDAC (500 kb upstream and 500 kb downstream), only the inflammatory bowel disease genomic regions remained borderline significantly associated with PDAC (*P*-value = 0.047). The top genes contributing to the inflammatory bowel disease association after excluding the regions around known PDAC GWAS signals were *ACTR2* and *LINC00339* (*P*-value < 0.001). Genomic regions for celiac disease were not associated with PDAC (*P*-value=0.31).

Conclusions: Our results from a large consortium provide modest support for the hypothesis that genomic regions surrounding germline risk variants for gastrointestinal inflammatory diseases (based on published GWAS loci) are associated with PDAC.

Thesis Readers:

Dr. Elizabeth A. Platz

Dr. Nilanjan Chatterjee

Acknowledgements

I would like to take this opportunity to show my gratitude to the numerous people who supported me throughout my Master's studies. I am especially grateful to my research advisor at the National Cancer Institute (NCI), Dr. Rachael Stolzenberg-Solomon, for providing valuable guidance at every stage of this research project. It would not have been impossible for me to have this opportunity to do my thesis work at NCI without her continued support and mentoring. I am also thankful for the support I received from other members of the Pancreatic Cancer Cohort Consortium (PanScan) and Pancreatic Cancer Case-Control Consortium (PanC4), including Dr. Kai Yu, Dr. Alison Klein and Dr. Rayjean J. Hung. In addition, I want to thank Dr. Han Zhang and Dr. William Wheeler for their help on statistical analysis.

I am deeply appreciative to my academic advisor at the Johns Hopkins Bloomberg School of Public Health, Dr. Elizabeth A. Platz, for her support throughout the program and advising on my research project. I am also thankful for the thesis readers, Dr. Elizabeth A. Platz and Dr. Nilanjan Chatterjee, for reading my thesis and offering valuable feedbacks.

Finally, and most importantly, I would like to thank my parents, grandparents, and other family members for their unconditional love and support throughout my childhood, adolescence, and adulthood. I know I could never have had the chance to study in the United States without them. My debt to them is beyond any measure.

Table of Contents

List of Tables

1. Introduction

Pancreatic cancer is widely known for its increasing annual incidence and poor prognosis. [1] As of 2018, it is estimated that 55,440 people will be diagnosed with pancreatic cancer and 44,330 people will die of this disease, making it the 11th most common cancer type and the 3rd leading cause of cancer-related mortality in the United States. [1,2] Pancreatic ductal adenocarcinoma (PDAC) is the most common type of pancreatic cancer. [3] Effective, population-level screenings are not currently available for early disease diagnosis and management. [4] In addition to modifiable risk factors, such as cigarette smoking, and overweight and obesity, genetic susceptibility also contributes to risk, accounting for 5-10% of incident pancreatic cancer cases. [5-9]

Chronic gastrointestinal inflammation plays an important role in PDAC. Inflammatory bowel disease (IBD) patients have a high prevalence of pancreatic abnormalities, specifically idiopathic chronic pancreatitis. [10-13]. Swedish registry-based studies have identified significant excess incidence of PDAC among patients with Crohn's disease and ulcerative colitis (together known as IBD), compared to the general Swedish population, albeit based on a relatively small PDAC case numbers. [14,15] Celiac disease, which is characterized by chronic inflammation in small intestine, has been variably associated with PDAC. [16-18] A limitation of these studies is that chronic inflammatory intestinal diseases may not be clinically diagnosed, especially for those with mild symptoms. Therefore, the diseases may be underestimated in the epidemiologic studies that rely upon self-reported questionnaire or review of medical records. Association studies have suggested a genetic component underlying both IBD and celiac disease. [19-26]

In this study, we examined the association between genomic regions (20 kb upstream and 20 kb downstream) surrounding germline variants identified from published genome-wide association studies (GWAS) of inflammatory bowel disease and celiac disease, and risk of PDAC among a total of 8384 PDAC cases and 11955 controls from the Pancreatic Cancer Cohort Consortium (PanScan) and Pancreatic Cancer Case-Control Consortium (PanC4). [5,6,8,9] Our aim was to use genomic regions surrounding genetic variants that characterize these gastrointestinal inflammatory diseases as proxies for disease exposures in relation to PDAC.

2. Methods

2.1. Study sample

Our study was based on 9038 PDAC cases and 12389 controls of European ancestry from four previously conducted GWASs in PanScan and PanC4. [5,6,8,9] The methods for each of the four GWASs have previously been described. [5,6,8,9] Participants with non-exocrine pancreatic tumors (histology types 8150, 8151, 8153, 8155 and 8240) were not included because the etiology of these cancer is thought to be different. To facilitate testing of interactions by study design (cohort verse case-control), we excluded 654 cases and 434 controls derived from PanScan III case series or case-control studies because PanScan III used previously genotyped controls from cohort studies. We only included participants of European genetic ancestry to avoid the potential confounding due to population stratification. In total, 8384 PDAC cases and 11955 controls were included in our final analytic data set.

All participants gave informed consent and all studies were approved by the Institutional Review Board of each participating institution and the National Cancer Institute's Special Studies Institutional Review Board.

2.2. GWAS summary statistics

Genotype imputation across the four study phases was based on the 1000G (Phase 3, v1) reference dataset. [27] Due to the large overlap of variants on the arrays (Illumina HumanHap550 Infinium II, Human 610-Quad) used for PanScan I and II, these studies were combined and jointly analyzed, whereas PanScan III (OmniExpress, Omni1M, Omni2.5M and Omni5M) and PanC4 (Illumina HumanOmniExpressExome-8v1) were each analyzed separately. For each GWAS phase, unconditional logistic

regression was used to calculate odds ratios (OR) and 95% confidence intervals (CI), adjusting for age (10-year categories, (⩽50, 51-60, 61-70, 71-80, and ⩾81)), sex, study, geographic region, and the top eigenvectors. Summary statistics for each of the phases were generated and used for analysis.

2.3. SNPs selection

We identified single-nucleotide polymorphisms (SNPs) associated with Crohn's disease, ulcerative colitis, IBD (Crohn's disease and ulcerative colitis combined) and celiac disease at genome-wide association level (P-value < 5×10^{-8}) using the most recent and largest studies curated by NHGRI-EBI Catalog of published GWAS (https://www.ebi.ac.uk/gwas/) as of June 9, 2017. [19-26,28-41] We also reviewed each of the original GWAS publications and added additional SNPs that met the significance level criteria. [26] For each disease group, we mapped each germline variant to its closest gene and extended the genomic region 20 kb upstream and 20 kb downstream. We included genomic regions (50 kb upstream and 50 kb downstream) surrounding germline variants that do not map to any surrounding gene.

2.4. Statistical analysis

We employed the summary based adaptive rank truncated product (sARTP) method to test the association of the combined genomic regions for Crohn's disease, ulcerative colitis, IBD and celiac disease with PDAC. The sARTP method combines SNP-level association statistics across SNPs in a gene or inflammatory gastrointestinal disease group. [42] The associations from up to five most significant SNPs in a gene were accumulated. To control for false positives, the sARTP method employed a resampling procedure to adjust for the size of genes (i.e., number of known SNPs in a gene) and the

size of genes in each disease group. The gene- and disease-level association *P*-values were estimated from the resampled null distribution using one hundred million resampling steps. A panel of 503 European subjects (CEU, TSI, FIN, GBR, IBS) in the 1000 Genomes Project (phase 3, v1) was used in sARTP to estimate the linkage disequilibrium (LD) between SNPs. To eliminate the impact of population stratification, the square root of the genomic inflation factors for each study phase (λ = 1.02, 1.02, 1.01 and 1.06 for PanC4, PanScan I and II case-control, PanScan I and II cohort, and PanScan III cohort, respectively) was adjusted to rescale the standard error of the estimated log odds ratio at each SNP. To identify disease group associations not driven by GWAS signals, sensitivity analysis was performed by excluding the 1000-kb genomic region (500 kb upstream and 500 kb downstream) flanking the 21 previously published GWAS PDAC risk signals that could be caused by LD. These included signals at 1p36.33 (*NOC2L*), 1q32.1 (*NR5A2*), 2p13.3 (*ETAA1*), 3q29 (*TP63*), 5p15.33 (*CLPTM1L-TERT*), 7p13 (*SUGCT)*, 8q21.11 (*HNF4G*), 8q24.21 (*MYC*), 7q32.3 (*LINC-PINT*), 9q34.2 (*ABO*), 13q12.2 (*PDX1*), 13q22.1 (*KLF5/KLF12*), 16q23.1 (*BCAR1*), 17q12 (HNF1B), 17q25.1 (*LINC00673*), 18q21.32 (*GRP*), and 22q12.1 (*ZNRF3*). [5,6,8,9,43]

3. Results

The proportion of cases diagnosed at ages older than 60 years was 0.74 overall. A larger proportion of cases from the cohort studies were diagnosed at ages > 60 years compared to those from the case-control studies (Table 1).

3.1. Disease group and PDAC

The genomic regions for IBD (*P*-value = 0.0018), ulcerative colitis (*P*-value = 0.0030) and Crohn's disease (*P*-value = 0.037) were significantly associated with PDAC risk (Table 2). Genomic regions for celiac disease were not associated with PDAC (P-value = 0.31). After excluding the 1000-kb genomic region around the previously identified PDAC GWAS signals, only the association for IBD remained nominally significant (*P*-value = 0.047). Ulcerative colitis and Crohn's disease were no longer significantly associated with the outcome (*P*-value = 0.11 and 0.079, respectively). In stratified analyses by study design, the associations of genomic regions for IBD and ulcerative colitis with PDAC were statistically significant (*P*-value = 0.0094 and 0.0055, respectively) in case-control studies, but not in the cohort studies (*P*-value = 0.44 and 0.84, respectively) (Supplemental Table 1). Neither Crohn's nor celiac disease were significantly associated with PDAC in cohort and case-control studies.

3.2. Gene and PDAC

We computed gene-level *P*-values for the 382 genes included in the study (Table 3). 27 genes for IBD, 15 genes for ulcerative colitis, 18 genes for Crohn's disease, and 4 genes for celiac disease had *P*-values < 0.05. The most significant genes contributing to each disease group included *NR5A2*, *ACTR2*, *LINC00339* and *TMEM8C* for IBD; *NR5A2*, *LINC00339* and *TMEM8C* for ulcerative colitis and *ACTR2* and *TMEM8C* for

Crohn's disease at a *P*-value < 0.001. After excluding previously identified PDAC GWAS loci, the most significant genes associated with each inflammatory disease group at a *P*-value < 0.001 were *ACTR2* and *LINC00339* for IBD; *LINC00339* for ulcerative colitis and *ACTR2* for Crohn's disease (Supplemental Table 2).

3.3. SNP and PDAC

The association between individual SNP and PDAC within each gene that were significant at a *P*-value < 0.001 for each gastrointestinal disease group is shown in Table 4. 34 SNPs located in 16 genes within the IBD pathway, 17 SNPs in 7 genes within the ulcerative colitis pathway, and 21 SNPs in 10 genes within the Crohn's disease pathway were significantly associated with PDAC at a *P*-value < 0.001. Among the SNPs, 1 (rs2816950) was located in the *NR5A2* gene, 3 (rs268871, rs56792073, rs72822420) in *ACTR2*, 4 (rs3820290, rs140356857, rs60189068, rs2501299) in *LINC00339* and 5 (rs3094325, rs9331726, rs9330459, rs13293068, rs9286382) in *TMEM8C*. The D´ and r^2 between the susceptible locus for IBD (rs2816958) and the susceptible locus for PDAC (rs2816950) in the *NR5A2* among European population were 0.06 and 0.001, respectively using Broadway Institute SNP Annotation and Proxy Search (http://archive.broadinstitute.org/mpg/snap/ldsearchpw.php). [44]

4. Discussion

This pathway-based analysis of GWAS data showed that germline variations in the vicinity of susceptibility loci for gastrointestinal inflammatory diseases were significantly associated with PDAC. This association was attributed to the combined effect of germline variants located in *NR5A2*, *ACTR2*, *LINC00339* and *TMEM8C*. Associations were more statistically significant in the case-control studies than in the cohort studies. Removal of previously identified PDAC-associated genetic variants resulted in the loss of significant association of the ulcerative colitis and Crohn's disease group genomic regions, but the association between IBD and PDAC became borderline significant. *TMEM8C* was not significantly associated with PDAC after previously identified GWAS PDAC risk signals were excluded. We did not see an association between celiac disease and PDAC.

The significant association between IBD, ulcerative colitis and Crohn's disease pathway and PDAC is mainly driven by *NR5A2*, in which susceptible loci for both IBD and PDAC are located. [8,26] NR5A2, also known as liver receptor homolog (LRH-1), is an orphan nuclear receptor that is predominantly expressed in tissues of endodermal origin in mammals, such as the liver, exocrine pancreas and intestine. [45] NR5A2 shows pleiotropic effects and participates in a wide range of cellular processes. [46,47] Contrary to the role of NR5A2 in reverse cholesterol transport and bile acid homeostasis in liver, [48-56] emerging evidence has revealed NR5A2 as a key player in the pathogenesis of IBD and PDAC. [47,57] NR5A2 regulates intestinal immunity and attenuates the inflammation response by upregulating the local production of glucocorticoids. [58] Coste et al. showed that haploinsufficency of *Nr5a2* not only made mice more susceptible to developing IBD,

but also reduced the synthesis of glucocorticoids in intestine. [59] Consistent with the finding in the experimental mouse model, this study also identified reduced expression of *NR5A2* in the colon tissues of IBD patients. [59] Elevated expression of *NR5A2* in pancreatic ductal epithelium is associated with aggressive growth and proliferation of cancer cells in pancreas. [57,60,61] Benod et al. observed that the *NR5A2* mRNA transcript was increased 30-fold in human pancreatic cancer cells as opposed to normal pancreatic ductal epithelial cells. [57] In the same study, small interfering RNA (siRNA)-mediated knockdown of *NR5A2* substantially inhibited proliferation of pancreatic cells, [57] suggesting that activated *NR5A2* could result in pancreatic tumorigenesis. This hypothesis is supported by the established mechanisms of NR5A2 in the regulation of cell cycle, maintenance of pluripotent stem cells, and reprogramming of somatic cells. [47,57,62-66] Another possible mechanism is implied in the protective effect of NR5A2 against stress-induced apoptosis in pancreatic β cells via upregulating glucocorticoids synthesis. [67] However, it is unclear whether similar anti-apoptotic protection will be observed in the exocrine pancreas.

Genetic variants within *ACTR2*, *LINC00339* and *TMEM8C* in proximity to the IBD risk loci also contributed to the association between IBD and PDAC. Actin-related protein 2 (ARP2) encoded by *ACTR2* in humans, is a major subunit of ARP2/3 complex. [68] Although the function of ARP2 is yet to be determined, ARP2/3 complex has been shown to regulates actin cytoskeleton dynamics by participating in the assembly, organization and recycling of actin filaments. [69-75] Common cellular functions performed by ARP2/3 complex include cell migration, adhesion, and vesicle trafficking. [68,71-73,75-90] Recent studies indicate that defective function of either ARP2/3 complex or its activator

Wiskott–Aldrich syndrome protein (WASP) is associated with IBD and other autoimmune adverse events. [71,91-96] On the other hand, excess of ARP2/3 complex may render cells additional capacity to invade surrounding tissues, [97] a property often seen in cancer metastasis. [98] Accumulating evidence supports that expression of the ARP2/3 complex, together with WASP, is significantly upregulated in multiple tumor types, including pancreas. [99-104] The function of LINC00339 and TMEM8C is poorly understood in humans, and their biological links to both IBD and PDAC remain to be elucidated. One interesting observation that should be discussed here is that in the sensitivity analysis, the association between *TMEM8C* and PDAC in IBD, ulcerative colitis and Crohn's disease group completely disappeared after removing 21 previously identified GWAS hits for PDAC. In contrast, the association of both *ACTR2* and *LINC00339* remained intact. One hypothesis is that SNPs in *TMEM8C* could be in LD with GWAS risk signals for PDAC. In fact, both *TMEM8C* and *ABO* are located on chromosome 9. [5] However, the five SNPs identified in *TMEM8C* are poorly correlated with the GWAS hit in *ABO* (http://archive.broadinstitute.org/mpg/snap/ldsearchpw.php). [44] It could also be that *NR5A2* and one of the PDAC risk genes might act sequentially in the same biological pathway, in which the risk gene is an upstream regulator of *TMEM8C*. Alternatively, it is possible that both *NR5A2* and the risk gene might be independently governed by the same biological process, which could be disturbed if the risk gene is removed.

Of note, the significant association of ulcerative colitis was lost after previously identified PDAC-associated genetic variants were removed. One of the known PDAC GWAS signal in the *NR5A2*, rs2816938, is highly correlated with the susceptible locus

for PDAC identified in this study, rs2816950. We suspect that removal of rs2816938 in the sensitivity analysis also eliminates the effect of rs2816950, resulting in non-significant association between ulcerative colitis and PDAC. We observed more significant associations for ulcerative colitis and IBD among the case-control studies compared to cohort studies (Supplemental table 1). We speculate that this is mostly likely due to the larger number of cases and controls from the case-control studies and power to observe associations. Alternatively, the differences might be due to biases related to the selection of either the cases (e.g. survival bias) or controls from the case-control studies.

A strength of our study is its large number of PDAC cases and controls with genomewide genetic data and the ability to control for several confounding factors. Our findings should be interpreted cautiously considering several limitations. We lack direct clinical examination and diagnosis of these diseases and information on important environmental factors (e.g. gluten exposure) that play a role in the etiology of these diseases. Our associations are based on genomic regions identified from GWAS studies associated with Crohn's, ulcerative colitis, and celiac disease, which provides the opportunity to validly identify susceptibility to chronic inflammatory intestinal diseases and PDAC risk, even for those with IBD or celiac disease who may not be diagnosed. We employ statistical-based method to account for the cumulative effects of multiple SNPs associated with each disease group to evaluate genetic susceptibility for PDAC. [42,105-107] This inter-SNP interaction commonly seen in complex traits is overlooked in GWAS where the association of single SNP is tested independently, [40,108,109] resulting in "missing heritability". [110] Nevertheless, one limitation of this pathway approach is that it only

includes genomewide significant SNPs into analysis. Given that heritability is often distributed over thousands of genetic variants with small effects in complex traits, [111,112] the present study may overlook a significant portion of heritability. A complementary approach to investigate genetic overlap across diseases is LD Score regression, [111,113] which take the effect of all SNPs into account, regardless of their genomewide significance level. Our study may not be generalizable to non-European ancestry populations.

In summary, this study employed sARTP method to assess genetic susceptibility to inflammatory intestinal diseases in relation to risk of PDAC. Our results provide modest support for the hypothesis that genomic regions flanking GWAS-identified germline variants for inflammatory bowel diseases but not celiac disease are associated with PDAC. These findings may help better understand the etiology of PDAC. Further investigations are warranted to replicate our findings and to examine such associations in populations of non-European ancestry.

5. References

1. Jemal A, Ward EM, Johnson CJ, et al. Annual Report to the Nation on the Status of Cancer, 1975-2014, Featuring Survival. *J Natl Cancer Inst.* 2017;109(9).

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* 2018;68(1):7-30.

3. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. *N Engl J Med.* 2014;371(11):1039-1049.

4. U.S. Preventive Services Task Force. Final Evidence Review: Pancreatic Cancer. January 2014; https://www.uspreventiveservicestaskforce.org/Page/Document/final-evidence-review80/pancreatic-cancer-screening. Accessed March 23, 2018.

5. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet.* 2009;41(9):986-990.

6. Childs EJ, Mocci E, Campa D, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nat Genet.* 2015;47(8):911-916.

7. Klein AP. Genetic susceptibility to pancreatic cancer. *Mol Carcinog.* 2012;51(1):14-24.

8. Petersen GM, Amundadottir L, Fuchs CS, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat Genet.* 2010;42(3):224-228.

9. Wolpin BM, Rizzato C, Kraft P, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat Genet.* 2014;46(9):994-1000.

10. Herrlinger KR, Stange EF. The pancreas and inflammatory bowel diseases. *Int J Pancreatol.* 2000;27(3):171-179.

11. Pitchumoni CS, Rubin A, Das K. Pancreatitis in inflammatory bowel diseases. *J Clin Gastroenterol.* 2010;44(4):246-253.

12. Ramos LR, Sachar DB, DiMaio CJ, Colombel JF, Torres J. Inflammatory Bowel Disease and Pancreatitis: A Review. *J Crohns Colitis.* 2016;10(1):95-104.

13. Yadav D, Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. *Gastroenterology.* 2013;144(6):1252-1261.

14. Hemminki K, Li X, Sundquist J, Sundquist K. Cancer risks in ulcerative colitis patients. *Int J Cancer.* 2008;123(6):1417-1421.

15. Hemminki K, Li X, Sundquist J, Sundquist K. Cancer risks in Crohn disease patients. *Ann Oncol.* 2009;20(3):574-580.

16. Askling J, Linet M, Gridley G, Halstensen TS, Ekstrom K, Ekbom A. Cancer incidence in a population-based cohort of individuals hospitalized with celiac disease or dermatitis herpetiformis. *Gastroenterology.* 2002;123(5):1428-1435.

17. Elfstrom P, Granath F, Ye W, Ludvigsson JF. Low risk of gastrointestinal cancer among patients with celiac disease, inflammation, or latent celiac disease. *Clin Gastroenterol Hepatol.* 2012;10(1):30-36.

18. Ilus T, Kaukinen K, Virta LJ, Pukkala E, Collin P. Incidence of malignancies in diagnosed celiac patients: a population-based estimate. *Am J Gastroenterol.* 2014;109(9):1471-1477.

19. Anderson CA, Boucher G, Lees CW, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet.* 2011;43(3):246-252.

20. Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet.* 2008;40(8):955-962.

21. Dubois PC, Trynka G, Franke L, et al. Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet.* 2010;42(4):295-302.

22. Franke A, Balschun T, Sina C, et al. Genome-wide association study for ulcerative colitis identifies risk loci at 7q22 and 22q13 (IL17REL). *Nat Genet.* 2010;42(4):292-294.

23. Franke A, McGovern DP, Barrett JC, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet.* 2010;42(12):1118-1125.

24. Hunt KA, Zhernakova A, Turner G, et al. Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet.* 2008;40(4):395-402.

25. Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature.* 2012;491(7422):119-124.

26. Liu JZ, van Sommeren S, Huang H, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet.* 2015;47(9):979-986.

27. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74.

28. Burdett T, Hall PN, Hastings E, et al. The NHGRI-EBI Catalog of published genome-wide association studies. https://www.ebi.ac.uk/gwas/. Accessed June 9, 2017.

29. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* 2017;45(D1):D896-D901.

30. Consortium UIG, Barrett JC, Lee JC, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat Genet.* 2009;41(12):1330-1334.

31. Ellinghaus D, Ellinghaus E, Nair RP, et al. Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *Am J Hum Genet.* 2012;90(4):636-647.

32. Franke A, Balschun T, Karlsen TH, et al. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat Genet.* 2008;40(11):1319-1323.

33. Julia A, Domenech E, Chaparro M, et al. A genome-wide association study identifies a novel locus at 6q22.1 associated with ulcerative colitis. *Hum Mol Genet.* 2014;23(25):6927-6934.

34. Julia A, Domenech E, Ricart E, et al. A genome-wide association study on a southern European population identifies a new Crohn's disease susceptibility locus at RBX1-EP300. *Gut.* 2013;62(10):1440-1445.

35. Kenny EE, Pe'er I, Karban A, et al. A genome-wide scan of Ashkenazi Jewish Crohn's disease suggests novel susceptibility loci. *PLoS Genet.* 2012;8(3):e1002559.

36. Ostrowski J, Paziewska A, Lazowska I, et al. Genetic architecture differences between pediatric and adult-onset inflammatory bowel diseases in the Polish population. *Sci Rep.* 2016;6:39831.

37. Parkes M, Barrett JC, Prescott NJ, et al. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet.* 2007;39(7):830-832.

38. Raelson JV, Little RD, Ruether A, et al. Genome-wide association study for Crohn's disease in the Quebec Founder Population identifies multiple validated disease loci. *Proc Natl Acad Sci U S A.* 2007;104(37):14747-14752.

39. Silverberg MS, Cho JH, Rioux JD, et al. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet.* 2009;41(2):216-220.

40. Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447(7145):661-678.

41. Garner C, Ahn R, Ding YC, et al. Genome-wide association study of celiac disease in North America confirms FRMD4B as new celiac locus. *PLoS One.* 2014;9(7):e101428.

42. Zhang H, Wheeler W, Hyland PL, et al. A Powerful Procedure for Pathway-Based Meta-analysis Using Summary Statistics Identifies 43 Pathways Associated with Type II Diabetes in European Populations. *PLoS Genet.* 2016;12(6):e1006122.

43. Zhang M, Wang Z, Obazee O, et al. Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21. *Oncotarget.* 2016;7(41):66328-66343.

44. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics.* 2008;24(24):2938-2939.

45. Fayard E, Auwerx J, Schoonjans K. LRH-1: an orphan nuclear receptor involved in development, metabolism and steroidogenesis. *Trends Cell Biol.* 2004;14(5):250-260.

46. Cobo I, Martinelli P, Flandez M, et al. Transcriptional regulation by NR5A2 links differentiation and inflammation in the pancreas. *Nature.* 2018;554(7693):533-537.

47. Fernandez-Marcos PJ, Auwerx J, Schoonjans K. Emerging actions of the nuclear receptor LRH-1 in the gut. *Biochim Biophys Acta.* 2011;1812(8):947-955.

48. Chen F, Ma L, Dawson PA, et al. Liver receptor homologue-1 mediates species- and cell line-specific bile acid-dependent negative feedback regulation of the apical sodium-dependent bile acid transporter. *J Biol Chem.* 2003;278(22):19909-19916.

49. Delerive P, Galardi CM, Bisi JE, Nicodeme E, Goodwin B. Identification of liver receptor homolog-1 as a novel regulator of apolipoprotein AI gene transcription. *Mol Endocrinol.* 2004;18(10):2378-2387.

50. Freeman LA, Kennedy A, Wu J, et al. The orphan nuclear receptor LRH-1 activates the ABCG5/ABCG8 intergenic promoter. *J Lipid Res.* 2004;45(7):1197-1206.

51. Goodwin B, Jones SA, Price RR, et al. A regulatory cascade of the nuclear receptors FXR, SHP-1, and LRH-1 represses bile acid biosynthesis. *Mol Cell.* 2000;6(3):517-526.

52. Lee YK, Schmidt DR, Cummins CL, et al. Liver receptor homolog-1 regulates bile acid homeostasis but is not essential for feedback regulation of bile acid synthesis. *Mol Endocrinol.* 2008;22(6):1345-1356.

53. Lu TT, Makishima M, Repa JJ, et al. Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. *Mol Cell.* 2000;6(3):507-515.

54. Luo Y, Liang CP, Tall AR. The orphan nuclear receptor LRH-1 potentiates the sterol-mediated induction of the human CETP gene by liver X receptor. *J Biol Chem.* 2001;276(27):24767-24773.

55. Mataki C, Magnier BC, Houten SM, et al. Compromised intestinal lipid absorption in mice with a liver-specific deficiency of liver receptor homolog 1. *Mol Cell Biol.* 2007;27(23):8330-8339.

56. Nitta M, Ku S, Brown C, Okamoto AY, Shan B. CPF: an orphan nuclear receptor that regulates liver-specific expression of the human cholesterol 7alpha-hydroxylase gene. *Proc Natl Acad Sci U S A.* 1999;96(12):6660-6665.

57. Benod C, Vinogradova MV, Jouravel N, Kim GE, Fletterick RJ, Sablin EP. Nuclear receptor liver receptor homologue 1 (LRH-1) regulates pancreatic cancer cell growth and proliferation. *Proc Natl Acad Sci U S A.* 2011;108(41):16927-16931.

58. Mueller M, Cima I, Noti M, et al. The nuclear receptor LRH-1 critically regulates extra-adrenal glucocorticoid synthesis in the intestine. *J Exp Med.* 2006;203(9):2057-2062.

59. Coste A, Dubuquoy L, Barnouin R, et al. LRH-1-mediated glucocorticoid synthesis in enterocytes protects against inflammatory bowel disease. *Proc Natl Acad Sci U S A.* 2007;104(32):13098-13103.

60. Lin Q, Aihara A, Chung W, et al. LRH1 promotes pancreatic cancer metastasis. *Cancer Lett.* 2014;350(1-2):15-24.

61. Lin Q, Aihara A, Chung W, et al. LRH1 as a driving factor in pancreatic cancer growth. *Cancer Lett.* 2014;345(1):85-90.

62. Botrugno OA, Fayard E, Annicotte JS, et al. Synergy between LRH-1 and beta-catenin induces G1 cyclin-mediated cell proliferation. *Mol Cell.* 2004;15(4):499-509.

63. Gu P, Goodwin B, Chung AC, et al. Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development. *Mol Cell Biol.* 2005;25(9):3492-3505.

64. Heng JC, Feng B, Han J, et al. The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell.* 2010;6(2):167-174.

65. Wagner RT, Xu X, Yi F, Merrill BJ, Cooney AJ. Canonical Wnt/beta-catenin regulation of liver receptor homolog-1 mediates pluripotency gene expression. *Stem Cells.* 2010;28(10):1794-1804.

66. Zhou Q, Chipperfield H, Melton DA, Wong WH. A gene regulatory network in mouse embryonic stem cells. *Proc Natl Acad Sci U S A.* 2007;104(42):16438-16443.

67. Baquie M, St-Onge L, Kerr-Conte J, et al. The liver receptor homolog-1 (LRH-1) is expressed in human islets and protects {beta}-cells against stress-induced apoptosis. *Hum Mol Genet.* 2011;20(14):2823-2833.

68. Welch MD, DePace AH, Verma S, Iwamatsu A, Mitchison TJ. The human Arp2/3 complex is composed of evolutionarily conserved subunits and is localized to cellular regions of dynamic actin filament assembly. *J Cell Biol.* 1997;138(2):375-384.

69. Bailly M, Macaluso F, Cammer M, Chan A, Segall JE, Condeelis JS. Relationship between Arp2/3 complex and the barbed ends of actin filaments at the leading edge of carcinoma cells after epidermal growth factor stimulation. *J Cell Biol.* 1999;145(2):331-345.

70. Cooper JA, Schafer DA. Control of actin assembly and disassembly at filament ends. *Curr Opin Cell Biol.* 2000;12(1):97-103.

71. Goley ED, Welch MD. The ARP2/3 complex: an actin nucleator comes of age. *Nat Rev Mol Cell Biol.* 2006;7(10):713-726.

72. Machesky LM, Reeves E, Wientjes F, et al. Mammalian actin-related protein 2/3 complex localizes to regions of lamellipodial protrusion and is composed of evolutionarily conserved proteins. *Biochem J.* 1997;328 (Pt 1):105-112.

73. Mullins RD, Heuser JA, Pollard TD. The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proc Natl Acad Sci U S A.* 1998;95(11):6181-6186.

74. Schafer DA, Welch MD, Machesky LM, Bridgman PC, Meyer SM, Cooper JA. Visualization and molecular analysis of actin assembly in living cells. *J Cell Biol.* 1998;143(7):1919-1930.

75. Svitkina TM, Borisy GG. Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *J Cell Biol.* 1999;145(5):1009-1026.

76. Benesch S, Polo S, Lai FP, et al. N-WASP deficiency impairs EGF internalization and actin assembly at clathrin-coated pits. *J Cell Sci.* 2005;118(Pt 14):3103-3115.

77. Castellano F, Le Clainche C, Patin D, Carlier MF, Chavrier P. A WASp-VASP complex regulates actin polymerization at the plasma membrane. *EMBO J.* 2001;20(20):5603-5614.

78. Coppolino MG, Krause M, Hagendorff P, et al. Evidence for a molecular complex consisting of Fyb/SLAP, SLP-76, Nck, VASP and WASP that links the actin cytoskeleton to Fcgamma receptor signalling during phagocytosis. *J Cell Sci.* 2001;114(Pt 23):4307-4318.

79. DeMali KA, Barlow CA, Burridge K. Recruitment of the Arp2/3 complex to vinculin: coupling membrane protrusion to matrix adhesion. *J Cell Biol.* 2002;159(5):881-891.

80. Gasman S, Chasserot-Golaz S, Malacombe M, Way M, Bader MF. Regulated exocytosis in neuroendocrine cells: a role for subplasmalemmal Cdc42/N-WASP-induced actin filaments. *Mol Biol Cell.* 2004;15(2):520-531.

81. Jonsdottir GA, Li R. Dynamics of yeast Myosin I: evidence for a possible role in scission of endocytic vesicles. *Curr Biol.* 2004;14(17):1604-1609.

82. Kaksonen M, Sun Y, Drubin DG. A pathway for association of receptors, adaptors, and actin during endocytic internalization. *Cell.* 2003;115(4):475-487.

83. Kaksonen M, Toret CP, Drubin DG. A modular design for the clathrin- and actin-mediated endocytosis machinery. *Cell.* 2005;123(2):305-320.

84. Kelleher JF, Atkinson SJ, Pollard TD. Sequences, structural models, and cellular localization of the actin-related proteins Arp2 and Arp3 from Acanthamoeba. *J Cell Biol.* 1995;131(2):385-397.

85. Kovacs EM, Goodwin M, Ali RG, Paterson AD, Yap AS. Cadherin-directed actin assembly: E-cadherin physically associates with the Arp2/3 complex to direct actin assembly in nascent adhesive contacts. *Curr Biol.* 2002;12(5):379-382.

86. May RC, Caron E, Hall A, Machesky LM. Involvement of the Arp2/3 complex in phagocytosis mediated by FcgammaR or CR3. *Nat Cell Biol.* 2000;2(4):246-248.

87. Moreau V, Galan JM, Devilliers G, Haguenauer-Tsapis R, Winsor B. The yeast actin-related protein Arp2p is required for the internalization step of endocytosis. *Mol Biol Cell.* 1997;8(7):1361-1375.

88. Schaerer-Brodbeck C, Riezman H. Functional interactions between the p35 subunit of the Arp2/3 complex and calmodulin in yeast. *Mol Biol Cell.* 2000;11(4):1113-1127.

89. Weiner OD, Servant G, Welch MD, Mitchison TJ, Sedat JW, Bourne HR. Spatial control of actin polymerization during neutrophil chemotaxis. *Nat Cell Biol.* 1999;1(2):75-81.

90. Zhou K, Sumigray KD, Lechler T. The Arp2/3 complex has essential roles in vesicle trafficking and transcytosis in the mammalian small intestine. *Mol Biol Cell.* 2015;26(11):1995-2004.

91. Kahr WH, Pluthero FG, Elkadri A, et al. Loss of the Arp2/3 complex component ARPC1B causes platelet abnormalities and predisposes to inflammatory disease. *Nat Commun.* 2017;8:14816.

92. Bosticardo M, Marangoni F, Aiuti A, Villa A, Grazia Roncarolo M. Recent advances in understanding the pathophysiology of Wiskott-Aldrich syndrome. *Blood.* 2009;113(25):6288-6295.

93. Massaad MJ, Ramesh N, Geha RS. Wiskott-Aldrich syndrome: a comprehensive review. *Ann N Y Acad Sci.* 2013;1285:26-43.

94. Rotty JD, Wu C, Bear JE. New insights into the regulation and cellular functions of the ARP2/3 complex. *Nat Rev Mol Cell Biol.* 2013;14(1):7-12.

95. Villa A, Notarangelo L, Macchi P, et al. X-linked thrombocytopenia and Wiskott-Aldrich syndrome are allelic diseases with mutations in the WASP gene. *Nat Genet.* 1995;9(4):414-417.

96. Zhu Q, Zhang M, Blaese RM, et al. The Wiskott-Aldrich syndrome and X-linked congenital thrombocytopenia are caused by mutations of the same gene. *Blood.* 1995;86(10):3797-3804.

97. Wang W, Goswami S, Sahai E, Wyckoff JB, Segall JE, Condeelis JS. Tumor cells caught in the act of invading: their strategy for enhanced cell motility. *Trends Cell Biol.* 2005;15(3):138-145.

98. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-674.

99. Otsubo T, Iwaya K, Mukai Y, et al. Involvement of Arp2/3 complex in the process of colorectal carcinogenesis. *Mod Pathol.* 2004;17(4):461-467.

100. Semba S, Iwaya K, Matsubayashi J, et al. Coexpression of actin-related protein 2 and Wiskott-Aldrich syndrome family verproline-homologous protein 2 in adenocarcinoma of the lung. *Clin Cancer Res.* 2006;12(8):2449-2454.

101. Wang W, Goswami S, Lapidus K, et al. Identification and testing of a gene expression signature of invasive carcinoma cells within primary mammary tumors. *Cancer Res.* 2004;64(23):8585-8594.

102. Yanagawa R, Furukawa Y, Tsunoda T, et al. Genome-wide screening of genes showing altered expression in liver metastases of human colorectal cancers by cDNA microarray. *Neoplasia.* 2001;3(5):395-401.

103. Laurila E, Savinainen K, Kuuselo R, Karhu R, Kallioniemi A. Characterization of the 7q21-q22 amplicon identifies ARPC1A, a subunit of the Arp2/3 complex, as a regulator of cell migration and invasion in pancreatic cancer. *Genes Chromosomes Cancer.* 2009;48(4):330-339.

104. Rauhala HE, Teppo S, Niemela S, Kallioniemi A. Silencing of the ARP2/3 complex disturbs pancreatic cancer cell migration. *Anticancer Res.* 2013;33(1):45-52.

105. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet.* 2010;11(12):843-854.

106. Yu K, Li Q, Bergen AW, et al. Pathway analysis by adaptive combination of P-values. *Genet Epidemiol.* 2009;33(8):700-709.

107. Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. *Am J Hum Genet.* 2007;81(6):1278-1283.

108. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet.* 2007;39(7):870-874.

109. Yeager M, Orr N, Hayes RB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet.* 2007;39(5):645-649.

110. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747-753.

111. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet.* 2015;47(11):1236-1241.

112. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet.* 2012;90(1):7-24.

113. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet.* 2015;47(3):291-295.

6. Tables

Table 1. Baseline Characteristics of Participants From PanScan and PanC4 Studies

Phase	Study design	Cases, n	Controls, n	Case diagnosis age > 60 (%)
PanScan I	Cohort	1383	1470	83.44
	Case-control	363	342	73.00
PanScan II	Case-control	1768	1841	66.06
PanScan III	Cohort	937	4651	93.92
PanC4	Case-control	3933	3651	69.70
Total	Cohort	2320	6121	87.67
	Case-control	6064	5834	68.83
	Combined	8384	11955	74.05

Table 2. Association [a] of Inflammatory Gastrointestinal Disease Groups and Risk of PDAC in PanScan and PanC4 Studies

Disease group	Gene, n	SNPs, n	*P*-value
Ulcerative colitis	163	13,522	0.0030
Crohn's disease	266	21,034	0.037
IBD	353	26,574	0.0018
Celiac disease	40	3,617	0.31
Excluding previous PDAC GWAS loci [b]			
Ulcerative colitis	156	12,925	0.11
Crohn's disease	262	20,521	0.079
IBD	344	25,860	0.047
Celiac disease	40	3,544	0.31

[a] The analysis was adjusted for age, sex, study, geographic region, and the top eigenvectors.

[b] Results were obtained after excluding 500kb region upstream and downstream of 21 PDAC GWAS risk signals.

Table 3. Association [a] of Inflammatory Gastrointestinal Disease Genes and Risk of PDAC at a $P < 0.05$ in PanScan and PanC4 Studies

Disease group	Gene	SNPs, n	*P*-value	Most significant SNP
IBD	*NR5A2*	129	1.0×10^{-7}	rs2816950
	ACTR2	30	2.3×10^{-4}	rs268871
	LINC00339	48	3.2×10^{-4}	rs3820290
	TMEM8C	99	6.0×10^{-4}	rs3094325
	RP11-594C13.1	147	0.0034	rs76145598
	rs1991866 [b]	59	0.0038	rs1372993
	AC007318.5	14	0.0071	rs6741232
	AZU1P1	20	0.0076	rs9548954
	COG6	83	0.0082	rs9532431
	NF2	50	0.012	rs9613968
	PPP3CA	129	0.014	rs756317
	CUX2	117	0.014	rs4766453
	F5	50	0.015	rs34067653
	FERMT1	107	0.016	rs11699483
	GS1-259H13.10	23	0.017	rs78680379
	ACER3	69	0.019	rs61902094
	RP11-213G21.2	43	0.019	rs80095830
	RP11-295B17.4	64	0.020	rs141601285
	RP11-295B17.2	55	0.022	rs60923396
	TMEM89	10	0.022	rs1264191
	TTC34	62	0.022	C1P2626472
	SLC26A6	11	0.023	rs1264191
	UQCRC1	15	0.026	rs1264191
	NCKIPSD	17	0.028	rs6763557
	ZNF114	167	0.029	rs58249067
	CELSR3	13	0.031	rs13096357
	BSN	27	0.041	rs9836291
Ulcerative colitis	*NR5A2*	172	1.0×10^{-7}	rs2816950
	TMEM8C	99	6.2×10^{-4}	rs3094325
	LINC00339	59	9.7×10^{-4}	rs3820290
	RP11-594C13.1	147	0.0034	rs76145598
	AZU1P1	20	0.0075	rs9548954
	GS1-259H13.10	23	0.017	rs78680379
	ACER3	69	0.019	rs61902094
	TMEM89	10	0.022	rs1264191
	RP11-295B17.4	67	0.022	rs141601285
	SLC26A6	11	0.023	rs1264191
	UQCRC1	15	0.025	rs1264191
	TTC34	81	0.026	C1P2626472

	BANK1	317	0.034	rs76831217
	ITGAL	33	0.041	rs4350585
	BSN	27	0.041	rs9836291
Crohn's disease	*ACTR2*	30	2.5×10^{-4}	rs268871
	TMEM8C	99	6.0×10^{-4}	rs3094325
	RP11-594C13.1	147	0.0034	rs76145598
	SFMBT1	73	0.0039	rs115632241
	RP11-894J14.5	73	0.0041	rs115632241
	rs9491697 [b]	24	0.0066	rs12525327
	AC007318.5	14	0.0071	rs6741232
	NF2	50	0.012	rs9613968
	PPP3CA	129	0.014	rs756317
	GS1-259H13.10	23	0.017	rs78680379
	ACER3	69	0.019	rs61902094
	RP11-295B17.2	55	0.022	rs60923396
	TMEM89	10	0.022	rs1264191
	SLC26A6	11	0.023d	rs1264191
	UQCRC1	15	0.025	rs1264191
	ZNF114	167	0.029	rs58249067
	BSN	27	0.041	rs9836291
	RP1-96H9.5	72	0.046	rs11064124
Celiac disease	*CUX2*	117	0.014446	rs4766453
	TTC34	86	0.027062	C1P2626472
	IL12A-AS1	201	0.031047	rs73171911
	COX17	39	0.03746	rs6785698

[a] The analysis was adjusted for age, sex, study, geographic region, and the top eigenvectors.

[b] SNP name is included as the gene name because the SNP does not map to any surrounding gene.

Table 4. Association [a] of Inflammatory Gastrointestinal Disease SNPs and Risk of PDAC at a $P < 0.001$ in PanScan and PanC4 Studies

Disease group	SNP	Gene	Chr.	Position	Alleles [b]	OR (95% CI)	*P*-value
IBD	rs2816950	*NR5A2*	1	199998491	G/C	0.82 (0.78-0.87)	2.4×10^{-13}
	rs268871	*ACTR2*	2	65481008	C/A	0.88 (0.82-0.93)	2.1×10^{-5}
	rs56792073	*ACTR2*	2	65512340	C/G	0.92 (0.88-0.96)	1.2×10^{-4}
	rs72822420	*ACTR2*	2	65513767	T/C	0.88 (0.82-0.94)	2.0×10^{-4}
	rs3820290	*LINC00339*	1	22333595	A/G	0.91 (0.87-0.96)	1.2×10^{-4}
	rs140356857	*LINC00339*	1	22334151	G/A	0.83 (0.75-0.92)	2.2×10^{-4}
	rs60189068	*LINC00339*	1	22336643	G/C	0.86 (0.80-0.93)	2.2×10^{-4}
	rs2501299	*LINC00339*	1	22345647	T/C	0.92 (0.88-0.96)	6.2×10^{-4}
	rs3094325	*TMEM8C*	9	136360877	A/G	0.78 (0.69-0.88)	3.7×10^{-5}
	rs9331726	*TMEM8C*	9	136368685	T/G	0.79 (0.70-0.89)	8.3×10^{-5}
	rs9330459	*TMEM8C*	9	136371224	T/G	0.79 (0.71-0.89)	8.7×10^{-5}
	rs13293068	*TMEM8C*	9	136379129	T/C	0.78 (0.69-0.89)	1.0×10^{-4}
	rs9286382	*TMEM8C*	9	136383216	T/C	0.78 (0.68-0.88)	1.3×10^{-4}
	rs76145598	*RP11-594C13.1*	14	87951925	A/G	0.73 (0.64-0.85)	1.9×10^{-5}
	rs183768453	*RP11-594C13.1*	14	87959580	A/C	0.75 (0.65-0.87)	1.6×10^{-4}
	rs117209303	*RP11-594C13.1*	14	87942082	G/T	0.76 (0.66-0.88)	2.6×10^{-4}
	rs1372993	*rs1991866*	8	129565662	G/A	0.86 (0.80-0.93)	4.2×10^{-5}
	rs16903111	*rs1991866*	8	129586776	T/C	0.88 (0.83-0.95)	3.7×10^{-4}
	rs6741232	*AC007318.5*	2	65424889	G/A	0.90 (0.85-0.95)	5.2×10^{-4}
	rs9548954	*AZU1P1*	13	40437892	A/G	0.77 (0.66-0.88)	2.0×10^{-4}
	rs9532431	*COG6*	13	40352443	G/A	0.77 (0.67-0.88)	1.1×10^{-4}
	rs117960821	*COG6*	13	40230066	C/T	0.79 (0.70-0.90)	2.6×10^{-4}
	rs9613968	*NF2*	22	29980315	C/T	0.92 (0.88-0.96)	1.4×10^{-4}
	rs756317	*PPP3CA*	4	102147410	T/C	0.90 (0.86-0.95)	2.1×10^{-4}
	rs2850958	*PPP3CA*	4	102171606	G/A	0.88 (0.82-0.94)	2.8×10^{-4}
	rs2119601	*PPP3CA*	4	102173886	C/T	0.89 (0.83-0.95)	8.3×10^{-4}
	rs4766453	*CUX2*	12	111692124	T/C	0.90 (0.86-0.95)	6.8×10^{-5}
	rs34067653	*F5*	1	169469188	T/C	0.92 (0.88-0.96)	2.1×10^{-4}
	rs11699483	*FERMT1*	20	6048460	G/A	0.88 (0.82-0.95)	5.7×10^{-4}
	rs62199232	*FERMT1*	20	6050137	T/G	0.88 (0.82-0.95)	7.5×10^{-4}
	rs1884649	*FERMT1*	20	6047155	G/A	0.88 (0.82-0.95)	7.5×10^{-4}
	C1P2626472	*TTC34*	1	2626472	C/A	0.82 (0.74-0.91)	1.9×10^{-4}
	rs58249067	*ZNF114*	19	48745291	C/T	0.88 (0.82-0.94)	1.4×10^{-4}
	rs7252339	*ZNF114*	19	48784141	A/C	0.92 (0.88-0.97)	8.3×10^{-4}
Ulcerative colitis	rs2816950	*NR5A2*	1	199998491	G/C	0.82 (0.78-0.87)	2.4×10^{-13}
	rs3094325	*TMEM8C*	9	136360877	A/G	0.78 (0.69-0.88)	3.7×10^{-5}
	rs9331726	*TMEM8C*	9	136368685	T/G	0.79 (0.70-0.89)	8.3×10^{-5}
	rs9330459	*TMEM8C*	9	136371224	T/G	0.79 (0.71-0.89)	8.7×10^{-5}
	rs13293068	*TMEM8C*	9	136379129	T/C	0.78 (0.69-0.89)	1.0×10^{-4}
	rs9286382	*TMEM8C*	9	136383216	T/C	0.78 (0.68-0.88)	1.3×10^{-4}

	rs3820290	*LINC00339*	1	22333595	A/G	0.91 (0.87-0.96)	1.2×10^{-4}
	rs140356857	*LINC00339*	1	22334151	G/A	0.83 (0.75-0.92)	2.2×10^{-4}
	rs60189068	*LINC00339*	1	22336643	G/C	0.86 (0.80-0.93)	2.2×10^{-4}
	rs2501299	*LINC00339*	1	22345647	T/C	0.92 (0.88-0.96)	6.2×10^{-4}
	rs76145598	*RP11-594C13.1*	14	87951925	A/G	0.73 (0.64-0.85)	1.9×10^{-5}
	rs183768453	*RP11-594C13.1*	14	87959580	A/C	0.75 (0.65-0.87)	1.6×10^{-4}
	rs117209303	*RP11-594C13.1*	14	87942082	G/T	0.76 (0.66-0.88)	2.6×10^{-4}
	rs9548954	*AZU1P1*	13	40437892	A/G	0.77 (0.66-0.88)	2.1×10^{-4}
	C1P2626472	*TTC34*	1	2626472	C/A	0.82 (0.74-0.91)	1.9×10^{-4}
	rs76831217	*BANK1*	4	102499219	G/A	0.90 (0.85-0.95)	2.1×10^{-4}
	rs2903269	*BANK1*	4	102450302	T/C	0.91 (0.86-0.96)	2.8×10^{-4}
Crohn's disease	rs268871	*ACTR2*	2	65481008	C/A	0.88 (0.82-0.93)	2.1×10^{-5}
	rs56792073	*ACTR2*	2	65512340	C/G	0.92 (0.88-0.96)	1.2×10^{-4}
	rs72822420	*ACTR2*	2	65513767	T/C	0.88 (0.82-0.94)	2.0×10^{-4}
	rs3094325	*TMEM8C*	9	136360877	A/G	0.78 (0.69-0.88)	3.7×10^{-5}
	rs9331726	*TMEM8C*	9	136368685	T/G	0.79 (0.70-0.89)	8.3×10^{-5}
	rs9330459	*TMEM8C*	9	136371224	T/G	0.79 (0.71-0.89)	8.7×10^{-5}
	rs13293068	*TMEM8C*	9	136379129	T/C	0.78 (0.69-0.89)	1.0×10^{-4}
	rs9286382	*TMEM8C*	9	136383216	T/C	0.78 (0.68-0.88)	1.3×10^{-4}
	rs76145598	*RP11-594C13.1*	14	87951925	A/G	0.73 (0.64-0.85)	1.9×10^{-5}
	rs183768453	*RP11-594C13.1*	14	87959580	A/C	0.75 (0.65-0.87)	1.6×10^{-4}
	rs117209303	*RP11-594C13.1*	14	87942082	G/T	0.76 (0.66-0.88)	2.6×10^{-4}
	rs115632241	*SFMBT1*	3	53040808	T/C	0.77 (0.68-0.87)	3.1×10^{-5}
	rs115632241	*RP11-894J14.5*	3	53040808	T/C	0.77 (0.68-0.87)	3.1×10^{-5}
	rs12525327	*rs9491697*	6	127107248	G/A	0.91 (0.86-0.95)	1.6×10^{-4}
	rs6741232	*AC007318.5*	2	65424889	G/A	0.90 (0.85-0.95)	5.2×10^{-4}
	rs9613968	*NF2*	22	29980315	C/T	0.92 (0.88-0.96)	1.4×10^{-4}
	rs756317	*PPP3CA*	4	102147410	T/C	0.90 (0.86-0.95)	2.1×10^{-4}
	rs2850958	*PPP3CA*	4	102171606	G/A	0.88 (0.82-0.94)	2.8×10^{-4}
	rs2119601	*PPP3CA*	4	102173886	C/T	0.89 (0.83-0.95)	8.3×10^{-4}
	rs58249067	*ZNF114*	19	48745291	C/T	0.88 (0.82-0.94)	1.4×10^{-4}
	rs7252339	*ZNF114*	19	48784141	A/C	0.92 (0.88-0.97)	8.3×10^{-4}
Celiac disease	rs4766453	*CUX2*	12	111692124	T/C	0.90 (0.86-0.95)	6.8×10^{-5}
	C1P2626472	*TTC34*	1	2626472	C/A	0.82 (0.74-0.91)	1.9×10^{-4}
	rs73171911	*IL12A-AS1*	3	159818194	G/A	0.82 (0.74-0.90)	8.6×10^{-5}

Abbreviations: Chr., chromosome; CI, confidence interval; OR, odds ratio.

[a] The analysis was adjusted for age, sex, study, geographic region, and the top eigenvectors.

[b] Reference allele/risk allele

Supplemental Table 1. Association [a] of Inflammatory Gastrointestinal Disease Group With Risk of PDAC by Study Design in PanScan and PanC4 Studies

Study design	Disease group	Gene, n	SNPs, n	*P*-value
Cohort [b]	Ulcerative colitis	162	13,353	0.84
	Crohn's disease	263	20,768	0.61
	IBD	350	26,259	0.44
	Celiac disease	40	3,592	0.44
Case-control [c]	Ulcerative colitis	163	13,513	0.0055
	Crohn's disease	266	20,926	0.60
	IBD	353	26,484	0.0094
	Celiac disease	40	3,601	0.15

[a] The analysis was adjusted for age, sex, study, geographic region, and the top eigenvectors.

[b] Cohort studies include PanScan I and PanScan III cohort studies.

[c] Case-control studies include PanScan I, PanScan II and PanC4 case-control studies.

Supplemental Table 2. Association [a] of Inflammatory Gastrointestinal Disease Gene and Risk of PDAC at a $P < 0.05$ After Excluding the 500kb Region Upstream and Downstream of the PDAC GWAS Risk Signals in PanScan and PanC4 Studies

Disease group	Gene	No. of SNPs	*P*-value	Most significant SNP
IBD	*ACTR2*	30	2.4×10^{-4}	rs268871
	LINC00339	48	3.3×10^{-4}	rs3820290
	RP11-594C13.1	147	0.0034	rs76145598
	AC007318.5	14	0.0070	rs6741232
	AZU1P1	20	0.0075	rs9548954
	COG6	83	0.0082	rs9532431
	NF2	50	0.012	rs9613968
	PPP3CA	129	0.014	rs756317
	CUX2	117	0.014	rs4766453
	F5	50	0.015	rs34067653
	FERMT1	107	0.016	rs11699483
	GS1-259H13.10	23	0.017	rs78680379
	ACER3	69	0.019	rs61902094
	RP11-213G21.2	43	0.019	rs80095830
	RP11-295B17.4	64	0.020	rs141601285
	RP11-295B17.2	55	0.022	rs60923396
	TMEM89	10	0.022	rs1264191
	TTC34	62	0.022	C1P2626472
	SLC26A6	11	0.023	rs1264191
	UQCRC1	15	0.025	rs1264191
	NCKIPSD	17	0.028	rs6763557
	ZNF114	167	0.029	rs58249067
	CELSR3	13	0.031	rs13096357
	BSN	27	0.041	rs9836291
Ulcerative colitis	*LINC00339*	59	9.8×10^{-4}	rs3820290
	RP11-594C13.1	147	0.0035	rs76145598
	AZU1P1	20	0.0075	rs9548954
	GS1-259H13.10	23	0.017	rs78680379
	ACER3	69	0.019	rs61902094
	RP11-295B17.4	67	0.022	rs141601285
	TMEM89	10	0.022	rs1264191
	SLC26A6	11	0.023	rs1264191
	UQCRC1	15	0.025	rs1264191
	TTC34	81	0.026	C1P2626472
	BANK1	317	0.034	rs76831217
	ITGAL	33	0.040	rs4350585
	BSN	27	0.041	rs9836291
Crohn's disease	*ACTR2*	30	2.4×10^{-4}	rs268871

	RP11-594C13.1	147	0.0034	rs76145598
	SFMBT1	73	0.0039	rs115632241
	RP11-894J14.5	73	0.0041	rs115632241
	rs9491697 [b]	24	0.0066	rs12525327
	AC007318.5	14	0.0071	rs6741232
	NF2	50	0.012	rs9613968
	PPP3CA	129	0.014	rs756317
	GS1-259H13.10	23	0.017	rs78680379
	ACER3	69	0.019	rs61902094
	RP11-295B17.2	55	0.022	rs60923396
	TMEM89	10	0.022	rs1264191
	SLC26A6	11	0.023	rs1264191
	UQCRC1	15	0.025	rs1264191
	ZNF114	167	0.029	rs58249067
	BSN	27	0.041	rs9836291
	RP1-96H9.5	72	0.046	rs11064124
Celiac disease	*CUX2*	117	0.014	rs4766453
	TTC34	86	0.027	C1P2626472
	IL12A-AS1	201	0.031	rs73171911
	COX17	39	0.038	rs6785698

[a] The analysis was adjusted for age, sex, study, geographic region, and the top eigenvectors.

[b] SNP name is included as the gene name because the SNP does not map to any surrounding gene.

CURRICULUM VITAE

FANGCHENG YUAN
3-604, 225 Shigu Road, Qinhuai
Nanjing, Jiangsu, China
(434) 466-6845
fyuan7@jhu.edu

EDUCATION

Master of Science in Epidemiology (Cancer) Expected May 2018
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Certificate in Clinical Trials Expected May 2018
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Bachelor of Science in Biochemistry (Distinction) May 2016
Bachelor of Arts in Biology May 2016
University of Virginia, Charlottesville, VA

RESEARCH EXPERIENCE

Master's Student Fellow June 2017-Present
Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD
• Examine genetic and metabolic risk factors for pancreatic cancer
• Summarize study results and prepare manuscript for publication

Research Assistant September 2017-Present
Johns Hopkins Surgery Center for Outcomes Research, Johns Hopkins University School of Medicine, Baltimore, MD
• Study surgical outcome improvement among pediatric patients
• Summarize study results and prepare manuscript for publication

Research Assistant January 2015-June 2016
Department of Cell Biology, University of Virginia School of Medicine, Charlottesville, VA
• Analyzed intracellular cholesterol regulation and insulin secretion
• Presented data at laboratory meeting and maintained laboratory notebook
• Assisted graduate students or laboratory technicians in training undergraduate students

Laboratory Technician January 2014-November 2014
University of Virginia International Genetically Engineered Machine Team, Charlottesville, VA
• Won Gold Medal Award at the 2014 Giant Jamboree International Conference in Boston, MA
• Constructed engineered *E. coli* bacteria and tested its effectiveness in microplastics degradation
• Presented data and research progress at laboratory meeting and maintained laboratory notebook
• Attended outreach activity to teach high school students basic biology laboratory techniques

VOLUNTEER EXPERIENCE

Acute Pediatrics Volunteer September 2015-April 2016
University of Virginia Children's Hospital, Charlottesville, VA

Emergency Department Volunteer September 2013-April 2015
University of Virginia Hospital, Charlottesville, VA

Health Unit Coordinator Assistant May 2013-July 2013
Transitional Care Hospital, Charlottesville, VA
• Visited patients and their families to provide companionship and support
• Assisted staff with paperwork, answering phones, restocking supplies, ordering wheelchairs, and running errands

TEACHING EXPERIENCE

Teaching Assistant August 2017-October 2017
Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
• Led a discussion section, held office hours to answer questions, and proctored course examinations
• Attended weekly instructors meeting to discuss course materials and logistics

Peer Teaching Assistant September 2015-December 2015
Department of Biology, University of Virginia College and Graduate School of Arts and Sciences, Charlottesville, VA
• Assisted graduate teaching assistant to distribute laboratory reagents, answer questions and clean up the laboratory area

POSTERS AND PUBLICATIONS

Yuan F, Zhang H, Hung RJ, Wheeler W, Platz EA, Amundadottir L, Jacobs E, Kraft P, Li D, Petersen G, Risch H, Wolpin B, Yu K, Klein A, Stolzenberg-Solomon R. Genetic susceptibility to chronic inflammatory intestinal diseases and pancreatic ductal adenocarcinoma: a pathway analysis of genome-wide association studies. Accepted for presentation at the 51st annual meeting of the Society for Epidemiologic Research, June 19-22, 2018.

Broshkevitch C, Inam H, Leehan J, Mantus G, Moss T, Schmitt A, Tucker M, Yao M, Yuan F, Ye Z. NYGONE: A wastewater biofilter for nylon microplastics. Presented at the annual International Genetically Engineered Machine Giant Jamboree, October 30-November 3, 2014.

HONORS AND AWARDS

Master's Tuition Scholarship, Johns Hopkins Bloomberg School of Public Health

Echols Scholar, University of Virginia

Intermediate Honors, University of Virginia

PROFESSIONAL DEVELOPMENT

Laboratory Skills: molecular cloning, site-directed mutagenesis, tissue culture, gene knockdown by RNA interference, Western blot, immunofluorescence, microscopy and spectrophotometry

Language Skills: Fluent in spoken and written Mandarin Chinese

Computer Skills: Microsoft Office Suite, STATA, REDCap, SQL

Training: CPR/First Aid Certified by American Red Cross (2016-Present)

Membership: Maryland Responds Medical Reserve Corps (2016-Present), Society for Epidemiologic Research (2017-Present)